

# Efficient algorithms for highly compressed data: The Word Problem in Higman's group is in P

Volker Diekert<sup>1</sup>, Jörn Laun<sup>1</sup>, Alexander Ushakov<sup>2</sup>

<sup>1</sup> FMI, Universität Stuttgart, Universitätsstr. 38, D-70569 Stuttgart, Germany

<sup>2</sup> Department of Mathematics, Stevens Institute of Technology, Hoboken, NJ 07030, USA

**Abstract.** Power circuits are data structures which support efficient algorithms for highly compressed integers. Using this new data structure it has been shown recently by Myasnikov, Ushakov and Won that the Word Problem of the one-relator Baumslag group is in P. Before that the best known upper bound has been non-elementary. In the present paper we provide new results for power circuits and we give new applications in algorithmic algebra and algorithmic group theory: 1. We define a modified reduction procedure on power circuits which runs in quadratic time thereby improving the known cubic time complexity. The improvement is crucial for our other results. 2. We improve the complexity of the Word Problem for the Baumslag group to cubic time thereby providing the first practical algorithm for that problem. 3. The main result is that the Word Problem of Higman's group is decidable in polynomial time. The situation for Higman's group is more complicated than for the Baumslag group and forced us to advance the theory of power circuits.

**Key words:** Data structures, Compression, Algorithmic group theory, Word Problem

## Introduction

*Power circuits* have been introduced in [20]. It is a data structure for integers which supports  $+$ ,  $-$ ,  $\leq$ , a restricted version of multiplication, and raising to the power of 2. Thus, by iteration it is possible to represent (huge) values involving the tower function by very small circuits. Another way to say this is that efficient algorithms for power circuits yield efficient algorithms for arithmetic with integers in highly compressed form. This idea of *efficient algorithms for highly compressed data* is the main underlying theme of the present paper. In this sense our paper is more about compression and data structures than about algorithmic group theory. However, the applications are in this area so far.

Indeed as a first application of power circuits, [21] showed that the Word Problem of the Baumslag group<sup>3</sup> is solvable in polynomial time. Algorithmic interests have a long history in combinatorial group theory. In 1910 Max Dehn [7] formulated fundamental algorithmic problems for (finitely presented) groups. The most prominent one is the *Word Problem*: "Given a finite presentation of some fixed group  $G$ , decide whether an input word  $w$  represents the trivial element  $1_G$  in  $G$ ." It took more than four decades until Novikov and Boone showed (independently) in the 1950's the existence of a fixed finitely presented group with an undecidable Word Problem, [22,3]. It is also true that there are finitely presented groups with a decidable Word Problem but

<sup>3</sup> Sometimes called Baumslag-Gersten group, e.g. in [23] or in a preliminary version of [21].

with arbitrarily high complexity, [24, Theorem 1.3]. In these examples the difficult instances for the Word Problem are extremely sparse, (because they encode Turing machine computations) and, inherently due to the constructions, these groups never appear in any natural setting.

In contrast, the Baumslag group  $G_{(1,2)}$  is given by a single defining relation, see Sect. 4. (It is a non-cyclic one-relator group all of whose finite factor groups are cyclic [1].) It has been a natural (and simplest) candidate for a group with a non-polynomial Word Problem in the worst case, because the Dehn function<sup>4</sup> of  $G_{(1,2)}$  is non-elementary by a result due to Gersten [10], see also [23]. Moreover, the only general way to solve the word problem in one-relator groups is by a Magnus break-down procedure [19,17] which computes normal forms. It was developed in the 1930s and there is no progress ever since. Its time-complexity on  $G_{(1,2)}$  is non-elementary, since it cannot be bounded by any tower of exponents.

So, the question of algorithmic hardness of the Word Problem in one-relator groups is still wide open. Some researchers conjecture it is polynomial (even quadratic, see [2]), based on observations on generic-case complexity [13]. Others conjecture that it cannot be polynomial, based on the fast growing Dehn functions. (Note that the Dehn function gives a lot information about the group. E.g., if it is linear, then the group is hyperbolic, and the Word Problem is linear. If it is computable, then the Word Problem is decidable [18].)

The contributions of the present paper are as follows: In a first part, we give new efficient manipulations of the data structure of *power circuits*. Concretely, we define a new reduction procedure called EXTENDTREE on power circuits. It improves the complexity of the reduction algorithm from cubic to quadratic time. This is our first result. It turns out to be essential, because reduction is a fundamental tool and applied as a black-box operation frequently. For example, with the help of a better reduction algorithm (and some other ideas) we can improve as our second result the complexity of the Word Problem in  $G_{(1,2)}$  significantly from  $\mathcal{O}(n^7)$  in [21] down to  $\mathcal{O}(n^3)$ , Thm. 3. This cubic algorithm is the first practical algorithm which works for that problem on all reasonably short instances.

The basic structure in our paper is the domain of rational numbers, where nominators are restricted to powers of two. Thus, we are working in the ring  $\mathbb{Z}[1/2]$ . We view  $\mathbb{Z}[1/2]$  as an Abelian group where multiplication with  $1/2$  is an automorphism. This in turn can be embedded into a semi-direct product  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$  which is the set of pairs  $(r, k) \in \mathbb{Z}[1/2] \times \mathbb{Z}$  with the multiplication  $(r, k)(s, \ell) = (r + 2^k s, k + \ell)$ . There is a natural partially defined swap operation which interchanges the first and second component. Semi-direct products (or more generally, wreath products) appear in various places as basic mathematical objects, and this makes the algebra  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$  with swapping interesting in its own right. This algebra has a Word Problem in  $\mathcal{O}(n^4)$  (Thm. 2). The Word Problem of  $G_{(1,2)}$  can be understood as a special case with the better  $\mathcal{O}(n^3)$  performance.

Another new application of power circuits shows that the Word Problem in Higman's group  $H_4$  is decidable in polynomial time. (We could also consider any  $H_q$  with  $q \geq 4$ .) This is our third and main result. Higman [12] constructed  $H_4$  in 1951 as the first example of a finitely presented infinite group where all finite quotient groups are trivial. This leads immediately to a finitely generated infinite simple quotient group of  $H_4$ ; and no such group was known before Higman's construction. The group  $H_4$  is constructed by a few operations involving amalgamation (see e.g. [26]). Hence, a Magnus break-down procedure (for amalgamated products) yields decidability of the Word Problem. The procedure computes normal forms, but the length of normal forms can be a tower function in the input length. (More accurately, one can show that the Dehn function of  $H_4$  has an order of magnitude as a tower function [4].) Thus, Higman's group has been another natural, but rather complicated candidate for a finitely presented group with an extremely hard

---

<sup>4</sup> We do not use any result about Dehn functions here, and we refer the interested reader e.g. to Wikipedia (or to the appendix) for a formal definition.

Word Problem. Our paper eliminates  $H_4$  as a candidate: We show that the Word Problem of  $H_4$  is in  $\mathcal{O}(n^6)$  (Thm. 4).

We obtain this result by new techniques for efficient manipulations of multiple markings in a single power circuit and their ability for huge compression rates. Compression techniques have been applied elsewhere for solving word problems, [14,15,25,11]. But in these papers the authors use straight-line programs whose compression rates are far too small (at best exponential) to cope with Baumslag or Higman groups.

Due to lack of space some few proofs are shifted to the appendix.

## 1 Notation and preliminaries

Algorithms and (decision) problems are classified by their *time complexity* on a random-access machine (RAM). Frequently we use the notion of *amortized analysis* with respect to a *potential function*, see e.g. in [6, Sect. 17.3].

The *tower function*  $\tau : \mathbb{N} \rightarrow 2^{\mathbb{N}}$  is defined as usual:  $\tau(0) = 1$  and  $\tau(i+1) = 2^{\tau(i)}$  for  $i \geq 0$ . Thus, e.g.  $\tau(4) = 2^{2^{2^{2^1}}} = 2^{16}$  and  $\tau(6)$  written in binary requires more bits than there are supposed to be electrons in this universe.

We use standard notation and facts from group theory as the reader can find in the classical text book [17]. In particular, we apply the standard (so called Magnus break-down) procedure for solving the word problem in HNN-extensions and amalgamated products. All HNN-extensions and amalgamated products in this paper have an explicit finite presentation. Amalgamated products are denoted by  $G *_A H$  where  $A$  is a subgroup in  $G$  and in  $H$ . The formal definition of  $G *_A H$  creates first a disjoint copy  $H'$  of  $H$ . Then one considers the free product  $G * H$  and adds defining relations identifying  $a \in A$  with its copy  $a' \in A'$ . We refer to [26] for the basic facts about Higman groups.

## 2 Power circuits

This section is based on [20], but we also provide new material like our treatment of multiple markings and improved time complexities.<sup>5</sup> Let  $\Gamma$  be a set and  $\delta$  be a mapping  $\delta : \Gamma \times \Gamma \rightarrow \{-1, 0, +1\}$ . This defines a directed graph  $(\Gamma, \Delta)$ , where  $\Gamma$  is the set of vertices and the set of directed arcs (or edges) is  $\Delta = \{(P, Q) \in \Gamma \times \Gamma \mid \delta(P, Q) \neq 0\}$  (the support of the mapping  $\delta$ ). Throughout we assume that  $(\Gamma, \Delta)$  is a *dag* (*directed acyclic graph*). In particular,  $\delta(P, P) = 0$  for all vertices  $P$ .

A *marking* is a mapping  $M : \Gamma \rightarrow \{-1, 0, +1\}$ . We can also think of a marking as a subset of  $\Gamma$  where each element in  $M$  has a sign (+ or -). (Thus, we also speak about a *signed subset*.) Each node  $P \in \Gamma$  is associated in a natural way with a marking, which is called its  $A$ -marking  $A_P$  and which is defined as follows:

$$A_P : \Gamma \rightarrow \{-1, 0, +1\}, \quad Q \mapsto \delta(P, Q)$$

Thus, the marking  $A_P$  is the signed subset which corresponds to the targets of outgoing arcs from  $P$ .

---

<sup>5</sup> In order to keep the paper self-contained and as we use a slightly different notation we give full proofs in the appendix.

We define the *evaluation*  $\varepsilon(P)$  of a node ( $\varepsilon(M)$  of a marking resp.) bottom-up in the dag by induction:

$$\begin{aligned}\varepsilon(P) &= 2^{\varepsilon(\Lambda_P)} && \text{for a node } P, \\ \varepsilon(M) &= \sum_P M(P)\varepsilon(P) && \text{for a marking } M.\end{aligned}$$

Note that leaves evaluate to 1, the evaluation of a marking is a real number, and the evaluation of a node  $P$  is a positive real number. Thus,  $\varepsilon(P)$  and  $\varepsilon(M)$  are well-defined. We have the following nice formula for nodes:  $\log_2(\varepsilon(P)) = \varepsilon(\Lambda_P)$ . Therefore we can view the marking  $\Lambda_P$  as "taking the logarithm of  $P$ ".

**Definition 1.** A power circuit is a pair  $\Pi = (\Gamma, \delta)$  with  $\delta : \Gamma \times \Gamma \rightarrow \{-1, 0, +1\}$  such that  $(\Gamma, \Delta)$  is a dag as above with the additional property that  $\varepsilon(M) \in \mathbb{Z}$  for all markings  $M$ .

We will see later in Cor. 2 that it is possible to check in quadratic time whether or not a dag  $(\Gamma, \Delta)$  is a power circuit. (One checks  $\varepsilon(\Lambda_P) \geq 0$  for all nodes  $P$ , Lem. 6 in the appendix.)

*Example 1.* We can represent every integer in the range  $[-n, n]$  as the evaluation of some marking in a power circuit with node set  $\{P_0, \dots, P_\ell\}$  such that  $\varepsilon(P_i) = 2^i$  for  $0 \leq i \leq \ell$  and  $\ell = \lfloor \log_2 n \rfloor$ . Thus, we can convert the binary notation of an integer  $n$  into a power circuit with  $\mathcal{O}(\log |n|)$  vertices and  $\mathcal{O}((\log |n|) \log \log |n|)$  arcs.

*Example 2.* A power circuit can realize tower functions, since a line of  $n + 1$  nodes allows to represent  $\tau(n)$  as the evaluation of the last node.

Sometimes it is convenient to think of a marking  $M$  as a formal sum  $M = \sum_P M(P)P$ . In particular,  $-M$  denotes a marking with  $\varepsilon(-M) = -\varepsilon(M)$ . For a marking  $M$  we denote by  $\sigma(M)$  its *support*, i.e.,

$$\sigma(M) = \{P \in \Gamma \mid M(P) \neq 0\} \subseteq \Gamma.$$

We say that  $M$  is *compact*, if we have  $\varepsilon(P) \neq \varepsilon(Q) \neq 2\varepsilon(P)$  for all  $P, Q \in \sigma(M)$ ,  $P \neq Q$ . If  $M$  is compact, then we have  $\varepsilon(M) = 0$  if and only if  $\sigma(M) = \emptyset$ , and we have  $\varepsilon(M) > 0$  if and only if  $M(P)$  is positive for the node  $P$  having the maximal value in  $\sigma(M)$ .

The insertion of a new node  $\text{CLONE}(P)$  without incoming arcs and with  $\Lambda_{\text{CLONE}(P)} = \Lambda_P$  is called *cloning of a node*  $P$ . It is extended to markings, where  $\text{CLONE}(M)$  is obtained by cloning all nodes in  $\sigma(M)$  and defining  $M(\text{CLONE}(P)) = M(P)$  for  $P \in \sigma(M)$  and  $M(\text{CLONE}(P)) = 0$  otherwise. We say that  $M$  is a *source*, if no node in  $\sigma(M)$  has any incoming arcs. Note that  $\text{CLONE}(M)$  is always a source.

If  $M = \sum_P M(P)P$  and  $K = \sum_P K(P)P$  are markings, then  $M + K = \sum_P (M(P) + K(P))P$  is a formal sum where coefficients  $-2$  and  $2$  may appear. For  $M(P) + K(P) = \pm 2$ , let  $P' = \text{CLONE}(P)$ . We define a marking  $(M + K)'$  by putting  $(M + K)'(P) = (M + K)(P) = \pm 1$ . In this way we can realize addition (and subtraction) in a power circuit by cloning at most  $|\sigma(M) \cap \sigma(K)|$  nodes.

Next, consider markings  $U$  and  $X$  with  $\varepsilon(U) = u$  and  $\varepsilon(X) = x$  such that  $u2^x \in \mathbb{Z}$  (e.g. due to  $x \geq 0$ ). We obtain a marking  $V$  with  $\varepsilon(V) = u2^x$  and  $|\sigma(V)| = |\sigma(U)|$  as follows. First, let  $V = \text{CLONE}(U)$  and  $X' = \text{CLONE}(X)$ . Next, introduce additional arcs between all  $P' \in \sigma(V)$  and  $Q' \in X'$  with  $\delta(P', Q') = X'(Q')$ . Note that the cloning of  $X$  avoids double arcs from  $V$  to  $X$ . The cloning of  $U$  is not necessary, if  $U$  happens to be a source.

We now introduce an alternative representation for power circuits which allows us to compare markings efficiently. The process of transforming a  $\Pi$  into this so-called tree representation is referred to as *reduction of a power circuit*.

**Definition 2.** A tree representation of a power circuit  $\Pi = (\Gamma, \delta)$  consists of

- i)  $\Gamma$  as a list  $[P_1, \dots, P_n]$  such that  $\varepsilon(P_i) < \varepsilon(P_{i+1})$  for all  $1 \leq i < n$ ,
- ii) a bit vector  $b(1), \dots, b(n-1)$  where  $b(i) = 1$  if and only if  $2\varepsilon(P_i) = \varepsilon(P_{i+1})$ , and
- iii) a ternary tree of height  $n$ , where each node has at most three outgoing edges, labeled by  $+1, 0$ , and  $-1$ . All leaves are at level  $n$ , and each leaf represents the marking  $M : \Gamma \rightarrow \{-1, 0, +1\}$  given by the labels of the unique path of length  $n$  from the leaf to the root of the tree. These markings must be compact. Furthermore, all markings  $\Lambda_P$  for  $P \in \Gamma$  are represented by leaves. Finally, for each level in the tree  $T$ , we keep a list of the nodes in that level.

If a path from some leaf to the root is labeled  $(0, +1, -1, 0, +1)$ , then that leaf represents the marking  $P_2 - P_3 + P_5$  and we know (due to compactness)  $\varepsilon(P_2) < 2\varepsilon(P_3) < 4\varepsilon(P_5)$ . The amount of memory for storing a tree representation is bounded by  $\mathcal{O}(|\Gamma| \cdot (\text{number of leaves}))$ . Part iii) of the definition is only needed inside the procedure EXTENDTREE, which is explained in the appendix. For simplicity the reader might want to ignore it in a first reading and just think of a tree representation as the power circuit graph plus an ordering of the nodes and a bit vector keeping track of doubles.

**Proposition 1 ([20]).** There is a  $\mathcal{O}(|\Gamma|)$  time algorithm which on input a tree representation  $\Delta$  of a power circuit and two markings  $K$  and  $M$  (given as leaves) compares  $\varepsilon(K)$  and  $\varepsilon(M)$ . It outputs whether the two values are equal and if not, which one of them is larger. In the latter case it also tells whether their difference is 1 or  $\geq 2$ .

*Proof.* Start at the root of the tree and go down the paths corresponding to  $K$  and  $M$  in parallel. The first pair of different labels on these paths determines the larger value. The check whether  $\varepsilon(K) = \varepsilon(M) + 1$  is equally easy due to compactness.  $\square$

**Definition 3.** Let  $\Pi = (\Gamma, \delta)$  be a power circuit. A chain (of length  $r$ ) in  $\Pi$  is a sequence of nodes  $(P_0, P_1, \dots, P_r)$  where  $\varepsilon(P_i) = 2^i \varepsilon(P_0)$  ( $0 \leq i \leq r$ ). A chain is maximal if it is not part of a longer chain. The number of maximal chains in  $\Pi$  is denoted  $c(\Pi)$ . We define the potential of  $\Pi$  to be  $\text{pot}(\Pi) = c(\Pi) \cdot |\Gamma|$ .

The following statement uses amortized time w.r.t. the potential function  $\text{pot}(\Pi)$ . Note that the potential  $\text{pot}(\Pi)$  remains bounded by  $|\Gamma|^2$  since  $c(\Pi) \leq |\Gamma|$ .

**Theorem 1.** The following procedure EXTENDTREE runs in amortized time  $\mathcal{O}((|\Gamma| + |U|) \cdot |U|)$ :

*Input:* A dag  $\Pi = (\Gamma \cup U, \delta)$ , where  $\Gamma$  and  $U$  are disjoint with no arcs pointing from  $\Gamma$  to  $U$  and such that  $(\Gamma, \delta|_{\Gamma \times \Gamma})$  is a power circuit in tree representation. The potential is defined by the potential of its power circuit-part  $\text{pot}((\Gamma, \delta|_{\Gamma \times \Gamma}))$ . The output of the procedure is "no", if  $\Pi$  is not a power circuit (because  $\varepsilon(P) \notin \mathbb{Z}$  for some node  $P$ ). In the other case, the output is a tree representation of a power circuit  $\Pi' = (\Gamma', \delta')$  where:

- i)  $\Gamma \subseteq \Gamma'$  and  $\delta|_{\Gamma \times \Gamma} = \delta'|_{\Gamma \times \Gamma}$ .
- ii)  $|\Gamma'| \leq |\Gamma| + 3|U| + (c(\Pi) - c(\Pi'))$
- iii) For all  $Q \in U$  there exists a node  $Q' \in \Gamma'$  with  $\varepsilon(Q) = \varepsilon(Q')$ .
- iv) For every marking  $M$  in  $\Pi$  there exists a marking  $M'$  in  $\Pi'$  with  $\varepsilon(M') = \varepsilon(M)$  and  $|\sigma(M')| \leq |\sigma(M)|$ .

For  $\sigma(M) \subseteq \Gamma$  we can choose  $M = M'$  by i). If some further markings  $M_1, M_2, \dots, M_m$  are part of the input (those where  $\sigma(M) \cap U \neq \emptyset$ ), we need additional amortized time  $\mathcal{O}(|\sigma(M_1)| + \dots + |\sigma(M_m)| + m \cdot (|\Gamma| + |U|))$  to find the corresponding  $M'_1, M'_2, \dots, M'_m$ .

**Corollary 1.** *There is a  $\mathcal{O}(|\Gamma|^2)$  time procedure MAKE TREE that given a (graph representation) of a power circuit  $\Pi = (\Gamma, \delta)$  computes a tree representation. The number of nodes at most triples.*

*Proof.* This is a special case of Thm. 1 when  $U = \Pi$ . □

**Corollary 2.** *The test whether a dag  $(\Gamma, \delta)$  defines a power circuit can be done in  $\mathcal{O}(|\Gamma|^2)$ .*

The efficiency of MAKE TREE is crucial for all our results. In particular, Cor. 1 improves the cubic time complexity of [20] for reduction to quadratic.

### 3 Arithmetic in the semi-direct product $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$

The basic data structure for this paper deals with the semi-direct product  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$ . Here  $\mathbb{Z}[1/2]$  denotes the ring of rational numbers with denominators in  $2^{\mathbb{N}}$ . Thus, an element in  $\mathbb{Z}[1/2]$  is a rational number  $r$  which can be written as  $r = u2^x$  with  $u, x \in \mathbb{Z}$ . We view  $\mathbb{Z}[1/2]$  as an abelian group with addition. Multiplication by 2 defines an automorphism of  $\mathbb{Z}[1/2]$ , and hence the semi-direct product  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$  becomes a (non-commutative) group where elements are pairs  $(r, m) \in \mathbb{Z}[1/2] \times \mathbb{Z}$  and with the following explicit formula for multiplication:

$$(r, m) \cdot (s, n) = (r + 2^m s, m + n)$$

The semi-direct product  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$  is also isomorphic to a group with two generators  $a$  and  $t$  and the defining relation  $tat^{-1} = a^2$ . This group is known as the Baumslag-Solitar group  $\mathbf{BS}(1, 2)$ . The isomorphism from  $\mathbf{BS}(1, 2)$  to  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$  maps  $a$  to  $(1, 0)$  and  $t$  to  $(0, 1)$ . This is a homomorphism due to  $(0, 1)(1, 0)(0, -1) = (2, 0)$ . It is straightforward to see that it is actually bijective.

We have  $(r, m)^{-1} = (-r2^{-m}, -m)$  in  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$ , and a sequence of  $s$  group operations may lead to exponentially large or exponentially small values in the first component. Binary representation can cope with these values so there is no real need for power circuits when dealing with the group operation, only.

We equip  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$  with a partially defined *swap operation*. For  $(r, m) \in \mathbb{Z} \times \mathbb{Z} \subseteq \mathbb{Z}[1/2] \rtimes \mathbb{Z}$  we define  $\text{swap}(r, m) = (m, r)$ . This looks innocent, but note that a sequence of  $2^{\mathcal{O}(n)}$  defined operations starting with  $(1, 0)$  may yield a pair  $(0, \tau(n))$  where  $\tau$  is the tower function. Indeed  $\text{swap}(1, 0) = (0, 1) = (0, \tau(0))$  and

$$\text{swap}((0, \tau(n))(1, 0)(0, -\tau(n)) = \text{swap}(\tau(n+1), 0) = (0, \tau(n+1)). \quad (1)$$

However, we will show in section 6:

**Theorem 2.** *The Word Problem of the algebra  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$  with swapping is decidable in  $\mathcal{O}(n^4)$ .*

We use triples to denote elements in  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$ . A triple  $[u, x, k]$  with  $u, x, k \in \mathbb{Z}$  and  $x \leq 0 \leq k$  denotes the pair  $(u2^x, k+x) \in \mathbb{Z}[1/2] \rtimes \mathbb{Z}$ . For each element in  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$  there are infinitely many corresponding triples. Using the generators  $a$  and  $t$  of  $\mathbf{BS}(1, 2)$  one can write:

$$\begin{aligned} [u, x, k] &= (u2^x, k+x) = (0, x)(u, k) \in \mathbb{Z}[1/2] \rtimes \mathbb{Z} \\ &= t^x a^u t^k \in \mathbf{BS}(1, 2) \text{ and} \\ [u, x, k] \cdot [v, y, \ell] &= [u2^{-y} + v2^k, x+y, k+\ell] \end{aligned}$$

In the following we use power circuits with *triple markings* for elements in  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$ . We consider  $T = [U, X, K]$ , where  $U, X, K$  are markings in a power circuit with  $\varepsilon(U) = u$  and  $\varepsilon(X) = x \leq 0 \leq \varepsilon(K) = k$ ; and we define  $\varepsilon(T) \in \mathbb{Z}[1/2] \rtimes \mathbb{Z}$  to be the triple  $\varepsilon(T) = [u, x, k] = (u2^x, x+k)$ .

## 4 Solving the Word Problem in the Baumslag group

The Baumslag group  $G_{(1,2)}$  is a one-relator group with two generators  $a$  and  $b$  and the defining relation  $a^{a^b} = a^2$ . (The notation  $g^h$  means conjugation, here  $g^h = hgh^{-1}$ . Hence  $a^{a^b} = bab^{-1}aba^{-1}b^{-1}$ .) The group  $G_{(1,2)}$  can be written as an HNN extension of  $\mathbf{BS}(1, 2) \simeq \mathbb{Z}[1/2] \rtimes \mathbb{Z}$  with *stable letter*  $b$ ; and  $\mathbf{BS}(1, 2)$  is an HNN extension of  $\mathbb{Z} \simeq \langle a \rangle$  with *stable letter*  $t$ :

$$\begin{aligned} \langle a, b \mid a^{a^b} = a^2 \rangle &\simeq \langle a, t, b \mid a^t = a^2, a^b = t \rangle \\ &\simeq \text{HNN}(\langle a, t \mid a^t = a^2 \rangle, b, \langle a \rangle \simeq \langle t \rangle) \\ &\simeq \text{HNN}(\text{HNN}(\langle a \rangle, t, \langle a \rangle \simeq \langle a^2 \rangle), b, \langle a \rangle \simeq \langle t \rangle) \end{aligned}$$

Before the work of Myasnikov, Ushakov and Won ([21])  $G_{(1,2)}$  had been a possible candidate for a one-relator group with an extremely hard (non-elementary) word problem in the worst case by the result of Gersten [10]. (Indeed, the tower function is visible as follows: Let  $T(0) = t$  and  $T(n+1) = bT(n)aT(n)^{-1}b^{-1}$ . Then  $T(n) = t^{\tau(n)}$  by a translation of Eq. 1.) The purpose of this section is to improve the  $\mathcal{O}(n^7)$  time-estimation of [21] to cubic time. Thm. 3 yields also the first practical algorithm to solve the Word Problem in the Baumslag group for a worst-case scenario<sup>6</sup>.

**Theorem 3.** *The Word Problem of the Baumslag group  $G_{(1,2)}$  is decidable in time  $\mathcal{O}(n^3)$ .*

*Proof.* We assume that the input is already in compressed form given by a sequence of letters  $b^{\pm 1}$  and pairwise disjoint power circuits each of them with a triple marking  $[U, X, K]$  representing an element in  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$ , which in turn encodes a word over  $a^{\pm 1}$ 's and  $t^{\pm 1}$ 's.

We use the following invariants:

- i)  $U, X, K$  have pairwise disjoint supports.
- ii)  $U$  is a source.
- iii) All incoming arcs to  $X \cup K$  have their origin in  $U$ .
- iv) Arcs from  $U$  to  $X$  have the opposite sign of the corresponding node-sign in  $X$ .

These are clearly satisfied in case we start with a sequence of  $a^{\pm 1}$ 's,  $t^{\pm 1}$ 's, and  $b^{\pm 1}$ 's. The formula  $[u, x, k] \cdot [v, y, \ell] = [u2^{-y} + v2^k, x+y, k+\ell]$  allows to multiply elements in  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$  without destroying the invariants or increasing the total number of nodes in the power circuits (the invariants make sure that cloning is not necessary). The total number of multiplications is bounded by  $n$ . Taking into account that there are at most  $n^2$  arcs, we are within the time bound  $\mathcal{O}(n^3)$ .

Now we perform from left-to-right Britton reductions, see [17]. In terms of group generators this means to replace factors  $ba^s b^{-1}$  by  $t^s$  and  $b^{-1}t^s b$  by  $a^s$ . Thus, if we see a subsequence  $b[u, x, k]b^{-1}$ , then we must check if  $x+k = 0$  and after that if  $u2^x \in \mathbb{Z}$ . If we see a subsequence  $b^{-1}[u, x, k]b$ , then we must check  $u = 0$ . In the positive case we swap, in the other case we do nothing. Let us give the details: For a test we compute a tree representation of the circuit using MAKE TREE which takes time  $\mathcal{O}(n^2)$ . After each test for a Britton reduction, the tree representation is deleted. There are two possibilities for necessary tests.

- 1.)  $u = 0$ . If yes, remove in the original power circuit the source  $U$ , this makes  $X \cup K$  a source; replace  $[u, x, k]$  by  $[x+k, 0, 0]$ . The invariants are satisfied.

<sup>6</sup> It is easy to design simple algorithms which perform extremely well on random inputs. But for all these algorithms fail on short instances, e.g. in showing  $tT(6) = T(6)t$ .



- 2.)  $x + k = 0$ . If yes, check whether  $u2^x \in \mathbb{Z}$ . If yes, replace  $[u, x, k]$  in the original power circuit by either  $[0, u2^x, 0]$  or  $[0, 0, u2^x]$  depending on whether  $u2^x$  is negative or positive. We get  $u2^x$  without increasing the number of nodes, since arcs from  $U$  to  $X$  have the opposite signs of the node-signs in  $X$ . Thus, if  $E$  has been the set of arcs before the test, it is switched to  $U \times X \setminus E$  after the test. The new marking for  $u2^x$  is a source and does not introduce any cycle, because its support is still the support of the source  $U$ .

It is easy to see that computing a Britton reduction on an input sequence of size  $n$ , we need at most  $2n$  tests and at most  $n$  of them are successful. Hence we are still within the time bound  $\mathcal{O}(n^3)$ .

At the end we have computed in time  $\mathcal{O}(n^3)$  a Britton-reduced normal form where inner parts (i.e. the ones not involving  $b^{\pm 1}$ 's) are given as disjoint power circuits. The result follows straightforwardly.  $\square$

## 5 Higman groups

The Higman group  $H_q$  has a finite presentation with generators  $a_1, \dots, a_q$  and defining relations  $a_p a_{p-1} a_p^{-1} = a_{p-1}^2$  for all  $p \in \mathbb{Z}/q\mathbb{Z}$ . From now on we interpret indices  $p$  for generators  $a_p$  as elements of  $\mathbb{Z}/q\mathbb{Z}$ . In particular,  $a_q = a_0$  and one of the defining relations says  $a_1 a_q a_1^{-1} = a_q^2$ . It is known [26] that  $H_q$  is trivial for  $q \leq 3$  and infinite for  $q \geq 4$ . Hence, in the following we assume  $q \geq 4$ . The group  $H_4$  was the first example of a finitely generated group where all finite quotient groups are trivial. It has been another potential natural candidate for a group with an extremely hard (non-elementary) word problem in the worst case. Indeed, define:

$$\begin{aligned} w(p, 0) &= a_p & \text{for } p \in \mathbb{Z}/q\mathbb{Z} \\ w(p-1, i+1) &= w(p, i) a_{p-1} w(p, i)^{-1} & \text{for } i \in \mathbb{N} \text{ and } p \in \mathbb{Z}/q\mathbb{Z} \end{aligned}$$

By induction,  $w(p, n) = a_p^{\tau(n)} \in H_q$ , where  $\tau(n)$  is the  $n$ -th value of the tower function, but the length of the words  $w(p, n)$  is  $2^{n+1} - 1$ , only. Hence there is a "tower-sized gap" between input length and length of a canonical normal form.<sup>7</sup>

For  $i, j \in \mathbb{N}$  with  $i \leq j$  we define the group  $G_{i, \dots, j}$  by the generators  $a_i, a_{i+1}, \dots, a_j \in \{a_1, \dots, a_q\}$ , and defining relations  $a_p a_{p-1} a_p^{-1} = a_{p-1}^2$  for all  $i < p \leq j$ . Note that each  $G_i \simeq \mathbb{Z}$  is the infinite cyclic group. The group  $G_{1, \dots, q}$  is not  $H_q$  because the relation  $a_1 a_q a_1^{-1} = a_q^2$  is missing, but  $H_q$  is a (proper) quotient of  $G_{1, \dots, q}$ . The groups  $G_{i, i+1}$  are, by the very definition, isomorphic to the Baumslag-Solitar group  $\mathbf{BS}(1, 2)$ , hence  $G_{i, i+1} \simeq \mathbb{Z}[1/2] \rtimes \mathbb{Z}$ . It is also clear that  $G_{i, \dots, j+1} \simeq G_{i, \dots, j} *_{G_j} G_{j, j+1}$  for  $j - i < q$ . Thus,  $G_{123} \simeq G_{12} *_{G_2} G_{2,3}$  and  $G_{341} \simeq G_{34} *_{G_4} G_{41}$ .

For simplicity we deal with  $q = 4$  only. The free group  $F_{13}$ , generated by  $a_1$  and  $a_3$  is a subgroup of  $G_{123}$  as well as a subgroup of  $G_{341}$ , see e.g. [26]. Thus we can build the amalgamated product  $G_{123} *_{F_{13}} G_{341}$  and a straightforward calculation shows

$$H_4 \simeq G_{123} *_{F_{13}} G_{341}.$$

This isomorphism yields a direct proof that  $H_4$  is an infinite group, see [26]. In the following we use the following well-known facts about amalgamated products, see [17, 26, 8]. The idea is to calculate an alternating sequence of group elements from  $G_{123}$  and  $G_{341}$ . The sequence can be shortened, only if one factor appears to be in the subgroup  $F_{13}$ . In this case we swap the factor

<sup>7</sup> This can be made more precise (and rigorous) by saying that the *Dehn-function* of  $H_4$  grows like a tower function ([4])



from  $G_{123}$  to  $G_{341}$  and vice versa. By abuse of language we call this procedure again a *Britton reduction*. (This is perhaps no standard notation in combinatorial group theory, but it conveniently unifies the same phenomenon in amalgamated products and HNN-extensions; and the notion of Britton reduction generalizes nicely to fundamental groups of graphs of groups.) Elements in the groups  $G_{i,i+1}$  are represented by triple markings  $T = [U, X, K]$  in some power circuit. In order to remember that we evaluate  $T$  in the group  $G_{i,i+1}$ , we give each  $T$  a *type*  $(i, i+1)$ , which is denoted as a subscript. For  $\varepsilon(T) = [u, x, k]$  we obtain:

$$\begin{aligned} \varepsilon(T_{(i,i+1)}) &= a_{i+1}^x a_i^u a_{i+1}^k && \in G_{i,i+1} \\ &= a_i^{u2^x} a_{i+1}^{x+k} && \text{if } u2^x \in \mathbb{Z} \end{aligned}$$

The following basic operations are defined for all indices  $i, i+1, i+2$  and  $i \in \mathbb{Z}/4\mathbb{Z}$ , but for better readability we just use indices 1, 2, and 3.

– Multiplication:

$$[u, x, k]_{(1,2)} \cdot [v, y, \ell]_{(1,2)} = [u2^{-y} + v2^k, x + y, k + \ell]_{(1,2)}. \quad (2)$$

– Swapping from (1, 2) to (2, 3):

$$[0, x, k]_{(1,2)} = [x + k, 0, 0]_{(2,3)}. \quad (3)$$

– Swapping from (2, 3) to (1, 2):

$$[z, 0, 0]_{(2,3)} = [0, 0, z]_{(1,2)} \text{ for } z \geq 0. \quad (4)$$

$$[z, 0, 0]_{(2,3)} = [0, z, 0]_{(1,2)} \text{ for } z < 0. \quad (5)$$

– Splitting:

$$[u, x, k]_{(1,2)} = [u2^x, 0, 0]_{(1,2)} \cdot [0, x, k]_{(1,2)} \text{ for } u2^x \in \mathbb{Z}. \quad (6)$$

$$[u, x, k]_{(2,3)} = [0, x, k]_{(2,3)} \cdot [u2^{-k}, 0, 0]_{(2,3)} \text{ for } u2^{-k} \in \mathbb{Z}. \quad (7)$$

From now on we work with a single power circuit  $\Pi$  together with a sequence  $T_j$  ( $j \in J$ ) of triple markings of various types. This is given as a tuple  $\mathcal{T} = (\Gamma, \delta; (T_j)_{j \in J})$ . We allow splitting operations only in combination with a multiplication, thus we never increase the number of triple markings inside  $\mathcal{T}$ . A tuple  $\mathcal{T} = (\Gamma, \delta; (T_j)_{j \in J})$ , where  $(\Gamma, \delta)$  is in tree representation is called a *main data structure*. We keep  $\mathcal{T}$  as a main data structure by doing addition and multiplication by powers of 2 using clones and calling EXTENDTREE on these after each basic operation.

**Definition 4.** The weight  $\omega(T)$  of a triple marking  $T = [U, X, K]$  is defined as

$$\omega(T) = |\sigma(U)| + |\sigma(X)| + |\sigma(K)|.$$

The weight  $\omega(\mathcal{T})$  of a main data structure  $\mathcal{T}$  is defined as  $\omega(\mathcal{T}) = \sum_{j \in J} \omega(T_j)$ .

Its size<sup>8</sup>  $\|\mathcal{T}\|$  is defined by  $\|\mathcal{T}\| = |\Gamma|$ .

**Proposition 2.** Let  $\mathcal{T} = (\Gamma, \delta; (T_j)_{j \in J})$  be a main data structure of size at most  $m$ , weight at most  $w$  (and with  $|J| + w \leq m$ ). The following assertions hold.

- i) No basic operation increases the weight of  $\mathcal{T}$ .
- ii) Each basic operation increases the size  $\|\mathcal{T}\|$  by  $\mathcal{O}(w + (c(\Pi) - c(\Pi)))$ .
- iii) Each basic operation takes amortized time  $\mathcal{O}(mw)$ .

<sup>8</sup> The definition is justified, since we ensure  $|J| + \omega(\mathcal{T}) \leq \|\mathcal{T}\|$  whenever arguing about  $\|\mathcal{T}\|$ .

iv) A sequence of  $s$  basic operations takes time  $\mathcal{O}(smw + m^2)$  and the size of  $\mathcal{T}$  remains bounded by  $\mathcal{O}(m + sw)$ .

*Proof.* Applying a basic operation means replacing the left-hand side of the equation by the right-hand side, thus forgetting any markings of the replaced triple(s). We can do the necessary tests, because we have a tree representation. For an operation we clone the involved markings, but this does not increase the weight. Note that there is time enough to create the clones with all their outgoing arcs. This yields the increase in the size by  $\mathcal{O}(w)$ . With the new clones we can perform the operations by using the algorithms described in section 2 on the graph representation of the circuit. We regain the main data structure by calling `EXTENDTREE` which integrates the modified clones into the tree representation.

In order to get iv) we observe that the initial number of maximal chains is  $m$  and there are at most  $\mathcal{O}(w)$  new ones created in each basic operation. Hence the total increase in size is  $\mathcal{O}(sw)$  and the difference in potential is at most  $m(m + sw)$ . The time bound follows.  $\square$

## 5.1 Solving the Word Problem in Higman's group

**Theorem 4.** *The Word Problem of  $H_4$  can be solved in time  $\mathcal{O}(n^6)$ .*

The rest of this section is devoted to the proof of Thm. 4. For solving the word problem in the Higman group  $H_4$  the traditional input is a word over generators  $a_p^{\pm 1}$ . We solve a slightly more general problem by assuming that the input consists of a single power circuit  $\Pi = (I, \delta)$  together with a sequence of  $s$  triple markings of various types. Each triple marking  $[U, X, K]_{(p,p+1)}$  corresponds to  $a_{p+1}^{\varepsilon(X)} a_p^{\varepsilon(U)} a_{p+1}^{\varepsilon(K)} \in H_4$ .

Let us fix  $w$  to be the total weight of  $\mathcal{T} = (I, \delta; (T_j)_{1 \leq j \leq s})$ . For simplicity we assume  $s \leq w$  and that  $w$  and sizes of clones are bounded by  $\|T\| = |I|$ . (This is actually not necessary, but it simplifies some bookkeeping.) Having  $s \leq w \leq n \in \mathcal{O}(w)$ , we can think of  $n = |I|$  as our input size. We transform the input  $\mathcal{T} = (I, \delta; (T_j)_{1 \leq j \leq s})$  into a main data structure by a call of `MAKETREE`.

During the procedure  $|I|$  increases, but the number of triple markings remains bounded by  $s$  and the weight remains bounded by  $w$ .

In order to achieve our main result we show how to solve the word problem with  $\mathcal{O}(s^2)$  basic operations on the main data structure  $\mathcal{T}$ . Assume we have shown this. Then, by Prop. 2, the final size will be bounded by  $m \in \mathcal{O}(s^2 w)$ ; and the time for all basic operations is therefore  $\mathcal{O}(s^4 w^2) \subseteq \mathcal{O}(n^6)$ .

We collect sequences of triple markings of type  $(1, 2)$  and  $(2, 3)$  in intervals  $\mathcal{L}$ , which in turn receive type  $(1, 2, 3)$ ; and we collect triple markings of type  $(3, 4)$  and  $(4, 1)$  in intervals of type  $(3, 4, 1)$ . Each interval has (as a sequence of triple markings) a semantics  $\varepsilon(\mathcal{L})$  which is a group element either in  $G_{123}$  or in  $G_{341}$  depending on the type of  $\mathcal{L}$ . Thus, it makes sense to ask whether  $\varepsilon(\mathcal{L}) \in F_{13}$ . These tests are crucial and dominate the runtime of the algorithm.

Now the sequence  $(T_j)_{1 \leq j \leq s}$  of triple markings appears as a sequence of intervals:

$$(\mathcal{L}_1, \dots, \mathcal{L}_f; \mathcal{L}_{f+1}, \dots, \mathcal{L}_t).$$

We introduce a separator ";" dividing the list in two parts.

The following invariants are kept up:

- i) All  $\mathcal{L}_1, \dots, \mathcal{L}_f$  satisfy  $\varepsilon(\mathcal{L}_i) \notin F_{13}$ . In particular, these intervals are not empty and they represent non-trivial group elements in  $(G_{123} \cup G_{341}) \setminus F_{13}$ .
- ii) The types of intervals left of the separator are alternating.

In the beginning each interval consists of exactly one triple marking, thus  $f = 0$  and  $t = s$ . The algorithm will stop either with  $1 \leq f = t$  or with  $f = 0$  and  $t = 1$ .

Now we describe how to move forward: Assume first  $f = 0$ . (Thus,  $t > 1$ .) If  $\varepsilon(\mathcal{L}_1) \notin F_{13}$ , then move the separator to the right, i.e. we obtain  $f = 1$ . If  $\varepsilon(\mathcal{L}_1) \in F_{13}$ , then, after possibly swapping  $\mathcal{L}_1$ , we join the intervals  $\mathcal{L}_1$  and  $\mathcal{L}_2$  into one new interval. In this case we still have  $f = 0$ , but  $t$  decreases by 1.

From now on we may assume that  $0 < f < t$ . If  $\mathcal{L}_f$  and  $\mathcal{L}_{f+1}$  have the same type, then append  $\mathcal{L}_{f+1}$  to  $\mathcal{L}_f$ , and move the separator to the left of  $\mathcal{L}_f$ . Thus, the values  $f$  and  $t$  decrease by 1.

If  $\mathcal{L}_f$  and  $\mathcal{L}_{f+1}$  have different types, then we test whether or not  $\varepsilon(\mathcal{L}_{f+1}) \in F_{13}$ . If  $\varepsilon(\mathcal{L}_{f+1}) \notin F_{13}$ , then move the separator to the right, i.e.  $t - f$  decreases by 1. If  $\varepsilon(\mathcal{L}_{f+1}) \in F_{13}$ , then we swap  $\mathcal{L}_{f+1}$  and join the intervals  $\mathcal{L}_f$  and  $\mathcal{L}_{f+1}$  into one new interval. Since we do not know whether the new interval belongs to  $F_{13}$ , we put the separator in front of it, decreasing both  $f$  and  $t$  by 1.

We have to give an interpretation of the output of this algorithm. Consider the case that we terminate with  $1 \leq f = t$ . Then  $\varepsilon(\mathcal{L}_1) \cdots \varepsilon(\mathcal{L}_t) \in H_4$  is a Britton-reduced sequence in the amalgamated product. It represents a non-trivial group element, because  $t \geq 1$ .

In the other case we terminate with  $f = 0$  and  $t = 1$ . We will make sure that the test " $\varepsilon(\mathcal{L}) \in F_{13}$ ?" can as a by-product also answer the question whether or not  $\varepsilon(\mathcal{L})$  is the trivial group element. If we do so, one more test on  $(\mathcal{L}_1)$  yields the answer we need.

Now, we analyze the time complexity. Termination is clear once we have explained how to implement a test " $\varepsilon(\mathcal{L}) \in F_{13}$ ?". Actually, it is obvious that the number of these tests is bounded by  $2s$ . Thus, it enough to prove the following claim.

**Lemma 1.** *Every test " $\varepsilon(\mathcal{L}) \in F_{13}$ ?" can be realized with  $\mathcal{O}(s)$  basic operations in the main data structure  $\mathcal{T}$ . The test yields either "no" or it says "yes" with the additional information whether or not  $\varepsilon(\mathcal{L})$  is the trivial group element. Moreover, in the "yes" case we can also swap the type of  $\mathcal{L}$  within the same bound on basic operations.*

*Proof.* Let us assume that  $\mathcal{L}$  is of type  $(1, 2, 3)$ , i.e., it contains only triples of types  $(1, 2)$  and  $(2, 3)$ . Let  $s$  be the length of  $\mathcal{L}$ . The group  $G_{123}$  is an amalgamated product where  $F_{13}$  is a free subgroup of rank 2, see [26] for a proof. In a first round we create a sequence of triple markings

$$(T_1, \dots, T_t)$$

with  $t \leq s$  such that for  $1 \leq i < t$  the type of  $T_i$  is  $(1, 2)$  if and only if the type of  $T_{i+1}$  is  $(2, 3)$ . We can do so by  $s - t$  basic multiplications from left-to-right without changing the semantics of  $g = \varepsilon(T_1) \cdots \varepsilon(T_t) \in G_{123}$ .

Next, we make this sequence Britton-reduced. Again, we scan from left to right. If we are at  $T = T_i$  with value  $[u, x, k]$  we have to check that either  $[u, x, k]_{(1,2)} = (0, z) \in \mathbb{Z}[1/2] \rtimes \mathbb{Z}$  or  $[u, x, k]_{(2,3)} = (z, 0) \in \mathbb{Z}[1/2] \rtimes \mathbb{Z}$  for some integer  $z \in \mathbb{Z}$ .

For the type  $(1, 2)$  we have  $[u, x, k]_{(1,2)} = (0, z)$  if and only if  $u = 0$ , which in a tree representation means that the support of the marking for  $u$  is empty. Hence this test is trivial. If the test is positive, we can replace  $[u, x, k]_{(1,2)}$  by  $[0, x, k]_{(1,2)}$  and we perform a swap to type  $(2, 3)$ . If  $t > 1$  we can recursively perform multiplications with its neighbors, thereby decreasing the value  $t$ .

For the type  $(2, 3)$  we have  $[u, x, k]_{(2,3)} = (z, 0)$  if and only if both  $k + x = 0$  and  $u2^x \in \mathbb{Z}$ . Tests are possible in linear time and if successful, we continue as in the precedent case.

The final steps are more subtle. Let  $\varepsilon(T_j) = g_j \in G_{12} \cup G_{23}$ . Recall that  $(g_1, \dots, g_t)$  is already a Britton-reduced sequence. We have  $g_1 \cdots g_t \in F_{13}$  if and only if there is a sequence  $(h_0, h_1, \dots, h_t)$  with the following properties:

- i)  $h_0 = h_t = 1$  and  $h_j \in G_2$  for all  $0 \leq j \leq t$ .
- ii)  $h_{j-1}g_j = g'_j h_j$  with  $g'_j \in G_1 \cup G_3$  for all  $1 \leq j \leq t$ .

Assume that such a sequence  $(h_0, h_1, \dots, h_t)$  exists. Then we have  $g'_j \in G_1$  if and only if  $g_j \in G_{12}$ . Moreover, whenever  $gh = g'h' \in G_{123}$  with  $g, g' \in G_1 \cup G_3$  and  $h, h' \in G_2$ , then  $g = g'$  and  $h = h'$ . This follows because  $g'^{-1}g = h'h^{-1} \in F_{13} \cap G_2 = \{1\}$ . Thus, the product  $h_{j-1}g_j$  uniquely defines  $g'_j \in G_1 \cup G_3$  and  $h_j \in G_2$ , because  $h_0 = 1$  is fixed.

The invariant during a computation from left to right is that  $\varepsilon(T_j) = h_{j-1}g_j$ . We obtain  $\varepsilon(T_j) = g'_j h_j$  by a basic splitting. If no splitting is possible we know that  $g \notin F_{13}$  and we can stop. If however a splitting is possible, then we have two cases. If  $j$  is the last index ( $j = t$ ), then, in addition, we must have  $h_j = 1$ . We can test this. If the test fails, we stop with  $g \notin F_{13}$ . If we are not at the last index we perform a swap. We split, then swap the right hand factor and multiply it with the next triple marking, which has the correct type to do so. As our sequence has been Britton-reduced the total number of triple markings remains constant. There can be no cancellations at this point. Thus, the test gives us the answer to " $\varepsilon(\mathcal{L}) \in F_{13}$ ?" using  $\mathcal{O}(s)$  basic operations. In the case  $\varepsilon(\mathcal{L}) \in F_{13}$  we still need to know whether  $\varepsilon(\mathcal{L}) = 1 \in G_{123}$ . For  $t > 1$  the answer is "no". It remains to deal with  $t = 1$ . But a test whether  $\varepsilon([u, x, k]) = 1$  just means to test both  $u = 0$  and  $x + k = 0$ .

Now, assume we obtain a "yes" answer and we know  $\varepsilon(\mathcal{L}) \in F_{13}$ . We do the swapping of types from left to right by using only the left factor in a splitting. These are additional  $s$  basic operations, hence the total number of  $\mathcal{O}(s)$  did not increase.  $\square$

## 6 Conclusion and future research

The Word Problem is a fundamental problem in algorithmic group theory. In some sense "almost all" finitely presented groups are hyperbolic and satisfy a "small cancelation" property, so the Word Problem is solvable in linear time! For hyperbolic groups there are also efficient parallel algorithms and the Word Problem is in  $\mathbf{NC}^2$ , see [5]. On the other hand, for many naturally defined groups little is known. Among one-relator groups the Baumslag group  $G_{(1,2)}$  was supposed to have the hardest Word Problem. But we have seen that it can be solved in cubic time. The method generalizes to the higher Baumslag groups  $G_{(m,n)}$  in case that  $m$  divides  $n$ , but this requires more "power circuit machinery" and has not worked out in full details yet, see [21]. The situation for  $G_{(2,3)}$  is open and related to questions in number theory. The Higman groups  $H_q$  belong to another family of naturally appearing groups where the Word Problem was expected to be non-polynomial. We have seen that the Word Problem in  $H_4$  is in  $\mathcal{O}(n^6)$ . It is easy to see that our methods show that the Word Problem in  $H_q$  is always in  $\mathbf{P}$ , but to date the exact time complexity has not been analyzed for  $q > 4$ .

Baumslag and Higman groups are built up via simple HNN extensions and amalgamated products. Many algorithmic problems are open for such constructions, for advances about theories of HNN-extensions and amalgamated products we refer to [16].

Another interesting open problem concerns the Word Problem in *Hydra* groups. Doubled hydra groups have Ackermannian Dehn functions [9], but still it is possible that their Word Problem is solvable in polynomial time.

## References

1. G. Baumslag. A non-cyclic one-relator group all of whose finite quotients are cyclic. *J. Austr. Math. Soc.*, 10(3-4):497–498, 1969.
2. G. Baumslag, A. G. Myasnikov, and V. Shpilrain. Open problems in combinatorial group theory. Second Edition. In *Combinatorial and geometric group theory*, volume 296 of *Contemporary Mathematics*, pages 1–38. American Mathematical Society, 2002.
3. W. W. Boone. The Word Problem. *Annals of Mathematics*, 70(2):207–265, 1959.
4. M. Bridson. Groups with quadratic Dehn functions. Unpublished. Lecture at "Geometry, Topology and Computation in Groups", Les Diablerets, Switzerland, 2010.
5. J.-Y. Cai. Parallel computation over hyperbolic groups. In *Proceedings 24th Annual ACM Symposium on Theory of Computing, STOC 92*, pages 106–115. ACM-press, 1992.
6. T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms*. The MIT Press, 3 edition, 2009.
7. M. Dehn. Über unendliche diskontinuierliche Gruppen. *Math. Ann.*, 71(1):116–144, 1911.
8. V. Diekert, A. J. Duncan, and A. G. Myasnikov. Geodesic rewriting systems and pregroups. In O. Bogopolski et al. editors, *Combinatorial and Geometric Group Theory*, Trends in Mathematics, pages 55–91. Birkhäuser, 2010.
9. W. Dison and T. R. Riley. Hydra groups. <http://arxiv.org/abs/1002.1945>, 2010
10. S. M. Gersten. Isodiametric and isoperimetric inequalities in group extensions. preprint, 1991.
11. N. Haubold and M. Lohrey. Compressed word problems in HNN-extensions and amalgamated products. In A. E. Frid et al. editors, *CSR*, volume 5675 of *LNCS*, pages 237–249. Springer, 2009.
12. G. Higman. A finitely generated infinite simple group. *London Math. Soc.*, 26:61–64, 1951.
13. I. Kapovich, A. G. Miasnikov, P. Schupp, and V. Shpilrain. Generic-case complexity, decision problems in group theory and random walks. *J. Algebra*, 264:665–694, 2003.
14. M. Lohrey. Word problems and membership problems on compressed words. *SIAM J. Comput.*, 35(5):1210–1240, 2006.
15. M. Lohrey and S. Schleimer. Efficient computation in groups via compression. In V. Diekert et al. editors, *CSR*, volume 4649 of *LNCS*, pages 249–258. Springer, 2007.
16. M. Lohrey and G. Sénizergues. Theories of HNN-extensions and amalgamated products. In M. Bugliesi et al. editors, *ICALP*, volume 4052 of *LNCS*, pages 504–515. Springer, 2006.
17. R. Lyndon and P. Schupp. *Combinatorial Group Theory*. Springer, 2001.
18. K. Madlener and F. Otto. Pseudo-natural algorithms for finitely generated presentations of monoids and groups. *J. Symb. Comput.*, 5:339–358, 1988.
19. W. Magnus. Das Identitätsproblem für Gruppen mit einer definierenden Relation. *Math. Ann.*, 106:295–307, 1932.
20. A. G. Myasnikov, A. Ushakov, and D. W. Won. Power Circuits, Exponential Algebra, and Time Complexity. Preprint at <http://arxiv.org/abs/1006.2570>, 2010.
21. A. G. Myasnikov, A. Ushakov, and D. W. Won. The Word Problem in the Baumslag group with a non-elementary Dehn function is polynomial time decidable. Preprint at <http://arxiv.org/abs/1102.2481>, 2011.
22. P. S. Novikov. On the algorithmic unsolvability of the word problem in group theory. *Trudy Mat. Inst. Steklov*, pages 1–143, 1955. In Russian.
23. A. N. Platonov. Isoparametric function of the Baumslag-Gersten group. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, pages 12–17, 2004. In Russian.
24. M. V. Sapir, J.-C. Birget, and E. Rips. Isoperimetric and Isodiametric Functions of Groups. *Ann. Math.*, 156:345–466, 2002.
25. S. Schleimer. Polynomial-time word problems. *Commentarii Mathematici Helvetici*, 83:741–765, 2008.
26. J.-P. Serre. *Trees*. New York, Springer, 1980.

## Appendix

### Reduction of power circuits

In this section we give a full proof of Thm. 1. We start with the observation that due to compactness of the markings in a tree representation, the leaves are automatically ordered by  $\varepsilon$ -value. The leftmost has the smallest and the rightmost has the largest  $\varepsilon$ -value. This easy calculation (essentially an argument about binary sums) is left to the reader.

Next, we establish some operations on tree representations.

**Lemma 2.** (*Insertion of a new node*) Let  $\Pi = (\Gamma, \delta)$  be a power circuit in tree representation and  $M$  a marking in  $\Pi$  given as a leaf in the tree. Then we can in amortized time  $\mathcal{O}(|\Gamma|)$  insert a new node  $P$  into the circuit with  $\Lambda_P = M$ .

*Proof.* If a node  $P$  with  $\varepsilon(P) = 2^{\varepsilon(M)}$  already exists, we abort immediately. Otherwise the index of the leaf corresponding to the (compact) marking  $M$  inside the list of leaves tells us the position of  $P$  in the sorting of  $\Gamma$ . (Actually, we have to count the number of leaves of type  $\Lambda_Q$  ( $Q \in \Gamma$ ) that are left of  $M$ .)

Next, we need to update the bit vector  $b$ . This is achieved by the procedure described in Prop. 1, which tells us whether  $\varepsilon(M) + 1 = \varepsilon(\Lambda_Q)$  where  $Q$  is the node succeeding  $P$  in the ordering of  $\Gamma$ .

Finally, we have to "stretch" the tree defined by  $\Delta$  by inserting a new level corresponding to the new node  $P$ . All the edges on that level have to be labeled by "0", as no marking uses the newly created node yet. This can be done in linear time using the lists of nodes we keep for each level.

Note that this may increase the potential by  $|\Gamma|$ , since both  $|\Gamma|$  and the number of chains might grow by one. This adds  $|\Gamma|$  to the amortized time, which is captured by the  $\mathcal{O}$ -notation.  $\square$

**Lemma 3.** (*Incrementation of a marking*) Given a marking  $M$  in a tree representation of  $\Pi = (\Gamma, \delta)$  one can generate in  $\mathcal{O}(|\sigma(M)|)$  time a marking  $M'$  with  $\varepsilon(M') = \varepsilon(M) + 1$ .

*Proof.* If  $\Gamma$  is empty, the claim is obvious. Otherwise let  $P$  be the unique node in  $\Gamma$  with  $\varepsilon(P) = 1$ . If  $M(P) \neq +1$ , we increment  $M(P)$  by one and are done. Otherwise, we look for a node  $P'$  with  $\varepsilon(P') = 2$ . If it doesn't exist (in which case  $\Gamma = \{P\}$  and  $\varepsilon(M) = 1$ ), we create it and put  $M(P) = 0$  and  $M(P') = +1$ . If  $P'$  does exist, then  $M(P') = 0$  due to compactness. Again, put  $M(P) = 0$  and  $M(P') = +1$ .

Note that in the last case the newly created marking is not necessarily compact and therefore cannot be inserted immediately into the tree representation as a leaf. We will deal with this in the next lemma.  $\square$

**Lemma 4.** (*Making a marking compact*) Let  $\Pi = (\Gamma, \delta)$  be a power circuit in tree representation and  $M$  be a marking in  $\Pi$  (not yet a leaf and in particular not compact; e.g. given as a list of signed pointer to nodes). Assume that for each node  $P \in \sigma(M)$  the last node  $T$  in the longest chain starting in  $P$  is not marked by  $M$ . Then  $M$  can be made compact in time  $\mathcal{O}(|\sigma(M)|)$  (and after that integrated into  $\Pi$  as a leaf in time  $\mathcal{O}(|\Gamma|)$ ) without changing the circuit and without increasing the weight of  $M$ .

*Proof.* We look at the nodes  $P \in \sigma(M)$  in ascending order (w.r.t. their value). There are essentially two ways for  $M$  not to be compact at a point  $P$ : Assume that  $M(P) = +1$  ( $-1$  is similar).

- 1.)  $P$  is the first node in a chain length 2 which  $M$  labels  $(+1, -1)$ . Replace it by  $(-1, 0)$ .
- 2.)  $P$  is the first node of a chain  $(P = P_1, P_2, \dots, P_k, P_{k+1})$  labeled  $M(P_i) = +1$  ( $1 \leq i \leq k$ ) and  $M(P_{k+1}) \neq +1$ . Note that by assumption  $P_{k+1}$  exists. Replace the labels by  $(-1, 0, \dots, 0, +1)$ . We might need to repeat this (if there is a node  $P_{k+2}$  with  $\varepsilon(P_{k+2}) = 2\varepsilon(P_{k+1})$  and  $M(P_{k+2}) = +1$ ) but this ultimately stops at  $T$ .

□

Often we need to increment a leaf marking by one and make it compact. In order to have the necessary nodes, we introduce the following concept:

**Definition 5.** Let  $\Pi = (\Gamma, \delta)$  be a power circuit in tree representation. A node  $J$  is called a joker, if it is the last in a maximal chain starting at the unique node with value 1 and  $J$  is not used in any leaf marking.

**Lemma 5.** Let  $\Pi = (\Gamma, \delta)$  be a power circuit in tree representation. Then we can in amortized time  $\mathcal{O}(|\Gamma|)$  insert a joker into the circuit. □

*Proof.* Start at the node  $P_0$  with  $\varepsilon(P_0) = 1$ . Using the bit vector, find the first "gap" in the chain starting at  $P_0$ , i.e., the largest  $i$  such that  $P_0, \dots, P_{n-1}$  exist with  $\varepsilon(P_i) = 2^i$ . Keep the number  $n$  in binary notation. Like in Lem. 4 compute a compact representation of  $n$ . Note that we are dealing with ordinary numbers here, not circuits! Use the compact representation of  $n$  to create  $P_n$  with  $\varepsilon(P_n) = 2^n$ . Check whether there is a node  $P_{n+1}$  and adjust the bit vector. If yes,  $P_n$  linked two maximal chains, so in amortized analysis we don't have to account for the  $\mathcal{O}(|\Gamma|)$  time used so far. Repeat the process until we create a node that is the end of a maximal chain. This is the joker. □

Now we are ready to prove Thm. 1.

*Proof of Thm. 1:* Let  $n = |\Gamma| + |U|$ . We may assume  $n \geq 1$ .

Perform a topological sorting of  $U$ , i.e. find an enumeration  $U = \{Q(0), \dots, Q(|U| - 1)\}$  such that there are no arcs from any  $Q(i)$  to  $Q(j)$  when  $i \leq j$ . Since  $\Pi$  is a dag, a topological ordering of  $U$  exists and it can be found in time  $\mathcal{O}(n|U|)$ , see e.g. [6]. The nodes of  $U$  will be moved to  $\Gamma$  in ascending topological order, so that all the time  $\Gamma$  remains a circuit, i.e., there are no arcs from  $\Gamma$  to  $U$ .

Let  $M$  be any one of the markings  $M_j$  ( $j = 1, \dots, m$ ) or  $\Lambda_P$  ( $P \in U$ ). While the nodes of  $U$  are being moved to  $\Gamma$ , there may be times when the support of  $M$  is partly in  $\Gamma$  and partly still in  $U$ . Later it will be completely contained in  $\Gamma$ . We will maintain the following invariants:

- i) Any marking whose support is completely contained in  $\Gamma$  is represented as a leaf (and thus compact).
- ii) For all other markings  $M$  and all nodes  $P \in \sigma(M) \cap \Gamma$  there is a chain starting at  $P$  and ending at a node  $T \notin \sigma(M)$  such that there is no node with double the value of  $\varepsilon(T)$  (i.e. the chain cannot be prolonged at the top end).

We now describe how to do the moving of the topologically smallest node  $Q$  of  $U$ .

We have  $\Lambda_Q \subseteq \Gamma$ , so by the invariant, this is a leaf. Hence it is possible to test whether  $\varepsilon(\Lambda_Q) < 0$  using Prop. 1. If this is the case,  $\Pi$  is not a power circuit and we stop with the output "no". From now on we assume  $\varepsilon(Q) \geq 1$ .

- 1.) *Insert a new joker*

See Lem. 5. Now we don't have to worry about incrementing compact markings anymore.

- 2.) *Find a replacement  $P$  for  $Q$  in  $\Gamma$*

Check whether there is a node  $P \in \Gamma$  with value  $\varepsilon(P) = \varepsilon(Q)$ . If not, use Lem. 2 to create it taking the marking  $\Lambda_Q$  as  $\Lambda_P$ . This takes amortized time  $\mathcal{O}(|\Gamma|)$ .



3.) *Adapt markings using  $Q$ :*

Using the bit vector, go up the chain in  $\Gamma$  starting at  $P$ . Prolong the chain at the top by creating a new node  $P'$  (use Lem. 3 on the successor marking of the last node of the chain, insert it into the tree via Lem. 4 and use it as a successor marking for creating the new node  $P'$  with Lem. 2. The time needed is  $\mathcal{O}(|\Gamma|)$ .

Go through all markings  $M$  ( $\Lambda_{Q'}$  for  $Q' \in U$  and  $M_j$  for  $j = 1, \dots, m$ ) that have  $Q \in \sigma(M)$ . Replace  $Q$  by  $P$  in  $M$ . If this leads to a double marking of  $P$  by  $M$ , replace those by the next node in the chain. Again, that node might become doubly marked by  $M$ , so repeat this. This stops at the latest at  $P'$  which is new and thus unmarked  $M$ . For each of these steps, the support of  $M$  decreased by one, so the total time (for all  $Q \in U$ ) is bounded by the sum of the sizes of all supports, i.e.,  $n \cdot |U|$  for successor markings and  $|\sigma(M_1)| + \dots + |\sigma(M_m)|$  for the markings  $M_j$  ( $j = 1, \dots, m$ ). Now  $Q$  is not part of any marking anymore and can be deleted.

4.) *Make markings compact:*

If  $Q$  was the last node of a marking  $M$  to be moved from  $U$  to  $\Gamma$ , we have to make  $M$  compact and create a leaf in the tree. This is done by using Lem. 4. Note that we have the invariant ii). We need time  $\mathcal{O}(|\sigma(M)| + n)$ , which over the whole procedure sums up to  $\mathcal{O}(n \cdot |U|)$  for successor markings and  $\mathcal{O}(|\sigma(M_1)| + \dots + |\sigma(M_m)| + m \cdot n)$  for the markings  $M_j$  ( $j = 1, \dots, m$ ).

5.) *Make room for later compactification:*

Start at  $P'$  and create a new node  $T$  that has value  $\varepsilon(T) = 2\varepsilon(P')$ . Check whether there is a node with value  $2\varepsilon(T)$ . Is yes, the creation of  $T$  has linked two maximal chains, thus decreasing the potential by  $|\Gamma|$ . This pays for the  $\mathcal{O}(|\Gamma|)$  time needed for creating  $T$  and the check. Repeat this until we create a  $T$  that has no node with double the value of  $T$ . Note that this takes only amortized time  $\mathcal{O}(|\Gamma|)$ .

□

## A more detailed look at power circuits

**Lemma 6.** *The following assertions are equivalent:*

- 1.)  $\varepsilon(P) \in 2^{\mathbb{N}}$  for all nodes  $P$ ,
- 2.)  $\varepsilon(\Lambda_P) \geq 0$  for all nodes  $P$ ,
- 3.)  $\varepsilon(M) \in \mathbb{Z}$  for all markings  $M$ .

*Proof.* Choose some node  $P$  without incoming arcs which exists because  $(\Gamma, \delta)$  defines a dag. The assertions are equivalent on  $\Gamma \setminus \{P\}$  by induction. The result now follows easily.

*Proof of Cor. 2:* The procedure MAKETREE uses EXTENDTREE which detects if  $\varepsilon(P) \notin 2^{\mathbb{N}}$ . The result follows from Lem. 6. □

## The Word Problem of the algebra $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$ with swapping

*Proof of Thm. 2:* The proof is almost the same as the proof of Thm. 3 which has been given above. Hence we focus on the differences in the proof. The input is given by a sequence of pairwise disjoint power circuits each of them with a triple marking  $[U, X, K]$  representing an element in  $\mathbb{Z}[1/2] \rtimes \mathbb{Z}$ . We use only the invariant that  $U, X, K$  have pairwise disjoint supports. Swapping of  $[u, x, k]$  is possible, if  $z = u2^x \in \mathbb{Z}$  and the result is either  $[x + k, z, 0]$  or  $[x + k, 0, z]$  depending on the sign of  $z$ . In order to realize a marking for  $z$  we clone  $U$ . This increases the size by  $|\sigma(U)|$ . At the end the size of the power circuit is quadratic in the input. This yields  $\mathcal{O}(n^4)$  time. □

## Dehn functions

No result about Dehn functions is used in our paper. However, for convenience of the interested reader we recall the definition of a Dehn function as given by Wikipedia. Let  $G$  be given by a finite generating set  $X$  with a finite defining set of relations  $R$ . Let  $F(X)$  be the free group with basis  $X$  and let  $w \in F(X)$  be a relation in  $G$ , that is, a freely-reduced word such that  $w = 1$  in  $G$ . Note that this is equivalent to saying that is,  $w$  belongs to the normal closure of  $R$  in  $F(X)$ . Hence we can write  $w$  as a sequence of  $m$  words  $xrx^{-1}$  with  $r \in R^{\pm 1}$  and  $x \in F(X)$ . The *area* of  $w$ , denoted  $\text{Area}(w)$ , is the smallest  $m \geq 0$  such that there exists such a representation for  $w$  as the product in  $F(X)$  of  $m$  conjugates of elements of  $R^{\pm 1}$ . Then the Dehn function of a finite presentation  $G = \langle X \mid R \rangle$  is defined as

$$\text{Dehn}(n) = \max \{ \text{Area}(w) \mid w = 1 \in G, |w| \leq n, \text{ and } w \text{ is freely-reduced} \}.$$

Two different finite presentations of the same group are equivalent with respect to *domination*. Consequently, for a finitely presented group the growth type of its Dehn function does not depend on the choice of a finite presentation for that group.

A function  $f : \mathbb{N} \rightarrow \mathbb{N}$  is dominated by  $g : \mathbb{N} \rightarrow \mathbb{N}$ , if there exists  $c \geq 1$  such that  $f(n) \leq cg(cn + c) + cn + c$  for all  $n \in \mathbb{N}$ .